



Řízená práce s digitálními dokumenty a jejich archivace

Milena TVRDÍKOVÁ, VŠB-TU Ostravaⁱ

Abstract

Amount of information that are necessary to process in our society is permanently growing. It implies the increasing number of digital documents and attention must be paid to their management and archiving. The paper deals with Content Management Systems, Document Management Systems and Enterprise Content Management. The aim of the paper is to provide a methodical guide for the implementation and innovation of such systems in companies/organizations and archiving of digital documents. The proposal is based on digital document flow and main criterion for selection of technology is minimization of risks given by components of proposed system.

Keywords

Content management, data archiving, document management system, enterprise content management, workflow.

JEL Classification: M15, M19

ⁱ Department of Applied informatics, Faculty of Economics, VŠB-Technical University of Ostrava, Sokolská 33, 701 21 Ostrava, Czech Republic.
milena.tvrdikova@vsb.cz

1. Úvod

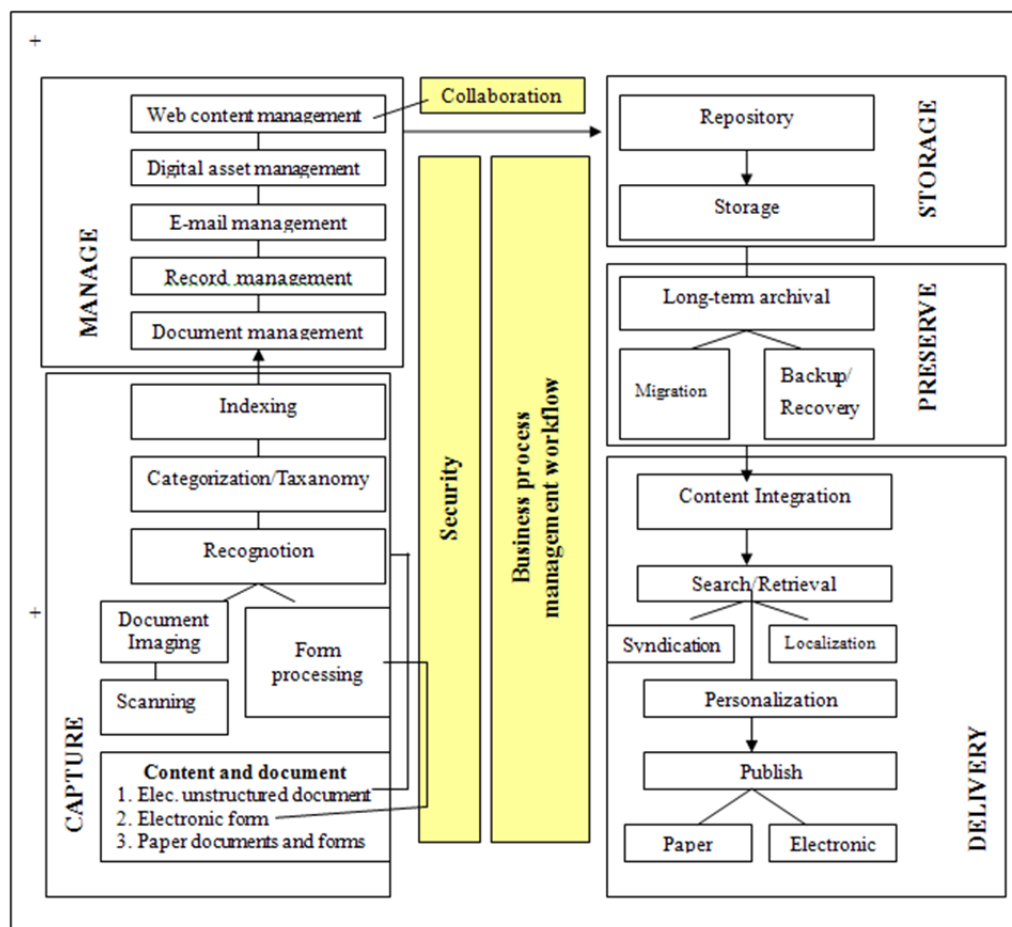
Množství zpracovávaných informací se neustále zvětšuje. Meziroční nárůst objemu informací se pohybuje v desítkách procent. Je tedy samozřejmé, že lidé stráví neúměrně mnoho času hledáním potřebné informace nebo potřebného dokumentu. Tím však nutně klesá produktivita a účinnost řízení a provádění administrativních činností. Práce s listinnými dokumenty je zdoluhavá.

Rozvoj informačních technologií umožňuje změnit charakter administrativních i řídicích činností. Automatizace přináší – přehlednost, dostupnost, sdílitelnost, aktuálnost, flexibilitu, zvýšení informovanosti a další výhody. Umožňuje také racionalizovat řadu administrativních činností a obchodních procesů zavedením workflow. V souvislosti s rozvojem informačních systémů a informačních technologií (dále IS/IT) a jejich aplikacemi v celospolečenském životě získávají tedy administrativní činnosti zcela nový charakter.

Cílem tohoto článku je vytvoření metodického rámce pro řešení přechodu na práci s digitálními dokumenty ve firmě nebo organizaci (nabídka charakteristik, uspořádání a zpřehlednění nabízených nástrojů a technologií, které práci s elektronickými dokumenty usnadňují a činí ji efektivnější). Tento metodický rámec by měl usnadnit orientaci řídicím pracovníkům při výběru jednotlivých komponent zaváděného či inovovaného systému pro práci s digitálními dokumenty. Pozornost je věnována taky archivaci digitálních dokumentů, včetně jejich dlouhodobé archivace. Návrh vychází z toku digitálních dokumentů a hlavním kritériem výběru doporučených technologií je minimalizace rizik spojených se specifickými vlastnostmi jednotlivých komponent vytvářeného systému.

2. Content management systémy

K tomu, abychom zvládali obrovské množství nejruznějších informací a dokumentů, nám pomáhají nástroje a technologie nazývané *Content management systems*. Chceme-li nějak pojem content vysvětlit,



Obrázek 1 Základní komponenty content management systému

Zdroj: Tětek, 2006

jeví se jako nevýstižnější jeho chápání jako dat s přesně definovanými charakteristikami (atributy) a metadaty pro vyhledávání. Systémy, o kterých budeme hovořit, umožňují vyhledání dokumentů na základě námi zadaných atributů, neumožňují však pouze vyhledávání, mají celou řadu dalších funkcí. Žádné komerčně dostupné řešení content management systému v současné době neobsahuje všechny dále popisované komponenty. Tato řešení jsou obvykle doplněna jiným komerčním řešením a řešení jsou spolu integrována.

Jak je patrné z obrázku 1, základní fáze zpracování digitálních dokumentů, odvozené ze životního cyklu dokumentu, jsou: získávání (*capture*), řízení (*manage*), spolupráce (*collaboration*), ukládání (*storage*), zabezpečení (*preserve*), doručení (*delivery*), řízení procesů a workflow (*business process management*) a bezpečnost (*security*), viz Tětek (2006).

2.1 Získávání (*capture*)

Na většině pracovišť se pracuje jak s listinnými, tak s elektronickými dokumenty, přičemž obojí jsou jak strukturované, tak nestrukturované.

Veškeré dokumenty můžeme rozdělit do tří základních typů:

- elektronické nestrukturované dokumenty (e-mail, texty a tabulky v elektronické podobě, digitální obrázky apod.), u kterých je nutné provést jejich kategorizaci a indexaci,
- elektronické formuláře (užívány jsou předdefinované šablony), kde je před kategorizací a indexací nutné provést rozpoznávání a zpracování formuláře,
- listinné dokumenty a formuláře, kde je nutné papírový dokument nejdříve převést do digitální podoby (*scanning*), následuje úprava do čitelné a zpracovatelné formy (*imaging*) a poté rozpoznávání, kategorizace a indexace.

Listinné dokumenty v informačním systému lze ponechat v dosavadní formě, ručně vést jejich

evidenci i informace o jejich uložení (např. došla korespondence, docházkové listy, žádanky o dovolenou, cestovní příkazy). Pak však zůstávají náklady na manipulaci s těmito dokumenty (kopírování, ukládání, vyhledávání, zabezpečení proti ztrátě atp.). Tento způsob nepodporují současné trendy v práci s dokumenty, např. legislativou podporovaná realizace datových schránek.

Druhou možností je přenos dat z dokumentů do informačního systému. Znamená to vedení elektronické evidence listinných dokumentů, tj. zápis základních údajů o dokumentu a přenos dat z dokumentu do informačního systému. Jelikož se většinou jedná o strukturované dokumenty, vkládají se data do připravených vstupních formulářů a pak se pracuje s digitální podobou dokumentu, listina je uložena do archivu. Snižuje se riziko ztráty dokumentu a zefektivní se proces jeho vyhledávání.

Třetí možností je skenování listinných dokumentů a jejich zpřístupnění v digitální podobě. Tento způsob snižuje časovou náročnost pořízení digitální formy dokumentu, minimalizuje se počet chyb vzniklých vkládáním dat, dokument je všem bezprostředně přístupný (přístup ošetřen přístupovými právy), minimalizuje se riziko ztráty dokumentu. Výrazně se snižují náklady na archivaci a skartaci. Klesají rovněž provozní náklady, protože se nepořizují kopie. Dosud tento způsob představoval vysoké prvotní náklady, avšak pozitivní vývoj cen hardware i software jej dělá stále atraktivnějším.

Firmám a institucím jsou nabízeny aplikace zajišťující přenos listinných dokumentů do informačního systému *Document Imaging Systems*, které umožňují:

- *Převod do elektronické podoby* (skenování), převod informací do grafické formy.
- *Zlepšování kvality skenovaných dokumentů* (*imaging*), provádí se např. určení prahové hodnoty černé a bílé (*thresholding*), zrušení natočení dokumentu (*deskew*), odstranění černých okrajů (*black border removal*), odstranění stínů (*deshade*), odstranění skvrn a teček (*despeckle*) a další funkce.
- *Rozpoznávání* znamená schopnost rozpoznání tištěných i ručně psaných znaků, kódů a značek a jejich převod bez manuálního přepisování dat (text nasnímaný skenerem nelze editovat a formátovat, k převodu z grafické do textové formy existují rozpoznávací programy).

Nejznámějšími technologiemi jsou:

- ICR (*Intelligent Character Recognition*) – určeno pro ručně psané písmo,
- OCR (*Optical Character Recognition*) – rozpozná tištěné a strojem psané písmo,
- BCR (*Bar Code Reading*) – převádí čárové kódy do podoby řetězců číslic a písmen,

- OMR (*Optical Mark Reading*) – převádí značky v podobě zaškrtnutých a zakřížkovaných okének na formuláři do digitální formy pro další zpracování.

- *Kategorizace* – automatické systematické zařazení dokumentu do content management system (dále CMS). Kategorizaci lze provést manuálně, hrozí však nebezpečí chybného zařazení (vliv lidského faktoru). Rozlišujeme:

- Typ dokumentu – složený (obsahuje více typů), zvukový záznam, obrazový záznam (statický nebo video záznam), text, datový soubor, počítačový program, interaktivní objekt.
- Formát dat obsažených v elektronickém dokumentu – textový, binární, otevřený, uzavřený.
- Úroveň zajištění autentičnosti – bez zajištění, digitální značka, digitální podpis, zaručený digitální podpis.
- Historická hodnota elektronického dokumentu – žádná, krátkodobá, střednědobá, dlouhodobá (Loebl, 2007).

- *Indexace a uložení* – integrace skenovaných dokumentů do informačního systému musí být ukončena jejich zaevidováním (typ dokumentu, datum vzniku, označení, ...) neboli indexací. Systém rozpozná typ dokumentu, podle identifikačních značek doplní další základní údaje a vše uloží do databáze, vytvářejí se tedy atributy (metadata) pro skenované dokumenty.

Význam metadat spočívá především ve vyhledávání elektronických dokumentů v archivu, v reprezentaci zobrazení elektronických dokumentů, v prokazování jejich autentičnosti a důvěryhodnosti a ve zjednodušení jejich správy.

Metadata pro elektronické dokumenty určené k uchování by měla obsahovat:

- popisná metadata (jednoznačný identifikátor dokumentu, popis obsahu),
- administrativní a technická metadata (popis formátu elektronického dokumentu, kontext elektronického dokumentu, historie elektronického dokumentu, ochrana elektronického dokumentu, řízení přístupu k elektronickému dokumentu),
- strukturální metadata (reference na externí zdroje).

Speciálně je metadatům věnována část kapitoly 5.2.

2.2 Správa dokumentů (*manage*)

Systémy pro efektivní správu dokumentů jsou realizovány pomocí jediného nebo kombinováním následující-

cích řešení. Jejich cílem je poskytnout okamžitý přístup ke správným dokumentům bez ohledu na jejich umístění a formát.

Document management systems (dále DMS) pomáhají lépe řídit vytváření, úpravu, schvalování a užívání digitálních dokumentů. Tyto systémy obsahují – vytváření knihoven dokumentů, profilování dokumentů, vyhledávání dokumentů, řízení verzí, revizi dokumentů a definování přístupových práv k dokumentu. Lze snadno zjistit, kde na síti a pod jakým jménem jsou soubory uloženy, kdo je jejich autorem, kolik různých verzí souboru existuje, kdo k nim má přístupová práva a jaká.

Digital asset management (správa multimediálních dat) má obdobné funkce jako DMS, je však zaměřen na *mediální dokumenty* (audio- nebo videozáznamy, fotografie).

Web content management (správa webových prezentací a aplikací) – je technologie zaměřená na vytváření, revizi, schvalování a publikování web dokumentů. Obsahuje nástroje pro vytváření, autorizaci a integraci dokumentů, nabídky prezentačních šablon a nástroje umožňující dynamické generování webových stránek.

2.3 Skupinová spolupráce (collaboration)

Potřeba vzájemné spolupráce při řešení jednotlivých úkolů vedla ke vzniku dalších softwarových produktů, které jsou známy pod pojmem groupware. Groupware je programové vybavení, které umožňuje lidem navzájem komunikovat, kooperovat na společném díle a koordinovat své aktivity.

Tyto aplikace usnadňují týmovou spolupráci, ale díky snadnému přístupu ke společným informacím a výkonným komunikačním prostředkům vedou často také k organizačním změnám. Podporují spolupráci lidí, kteří jsou geograficky rozptýleni, vedou ke zploštění hierarchické organizační struktury, k delegování pravomocí na pracovní týmy a zvyšují produktivitu práce těchto týmů. Z pohledu IS/IT může být spolupráce podporována samostatnými programovými produkty nebo komplexním groupwarovým systémem (který podporuje všechny tři formy spolupráce), viz Carda a Kunstová (2003).

2.4 Ukládání dokumentů (storage)

Ukládání dokumentů je další důležitou oblastí CMS. Nabízí se možnost ukládat dokumenty jako soubory na disk, což však může vyhovovat, pracujeme-li řádově s tisíci dokumenty. Pracujeme-li však s řádově větším počtem dokumentů, není to dostačující.

Většina dnešních CMS pracuje s tzv. dvouvrstvou technologií. První vrstvou je *archiv* (repository). Je nadefinována řada archivů a do jednotlivých archivů ukládáme dokumenty stejného typu. Archiv je jakýmsi

logickým uspořádáním dokumentů. Druhou vrstvou je vrstva fyzická (storage), *fyzické úložiště dokumentů* (např. DVD nebo disk). Výhodou dvouvrstvé technologie je skutečnost, že stále tentýž archiv může měnit své fyzické úložiště podle aktuálních potřeb a požadavků, ale z pohledu uživatele zůstávají repository neměnné.

2.5 Zabezpečení dokumentů (preserve)

Toto zabezpečení se netýká ochrany před zneužitím nebo neoprávněným přístupem k dokumentu. Jde o zajištění stavu, kdy dokumenty jsou stále k dispozici a vždy, když je to nutné, můžeme je zobrazit a užívat.

Tento požadavek vypadá samozřejmě, přihlédneme-li však k tempu vývoje informačních a komunikačních technologií, je třeba zabývat se:

- *dlohodobou archivaci* – např. uložení dokumentů vyžadujících dlouhodobé uložení na filmové médium;
- *migraci* – vypracování migračního plánu, v němž je definováno, kdy budou stará média nahrazena aktuální technologií, aby dokumenty byly stále přístupné, a také úroveň podpory dostupnosti dat dodavateli médií;
- *zálohováním* – stanovení zálohovací strategie a její dodržování.

2.6 Integrace a doručení dokumentů (delivery)

Integrace – každý informační systém určité firmy či organizace by měl mít vazby na veškeré aplikace, které jsou jeho součástí, tak aby došlo k vzájemnému propojování informací.

Při *doručení* můžeme dokument nalezený v CMS zpracovat různými způsoby vzhledem k jeho příjemci:

- *propojení* – dokument v našem CMS je upraven tak, aby jej bylo možno automaticky integrovat do jiného CMS;
- *personalizace* – dokument je upraven podle požadavků jednotlivých příjemců nebo skupin příjemců;
- *lokalizace* – dokument je upraven tak, aby splňoval legislativní a kulturní požadavky té země, pro kterou je určen.

2.7 Automatizace firemních procesů (workflow)

Workflow je automatizace celého nebo části podnikového procesu, během kterého jsou dokumenty, informace nebo úkoly předávány od jednoho účastníka procesu ke druhému podle sady procedurálních pravidel, viz Tvrdíková (2008b).

Workflow systém je tedy nástroj umožňující automatizovat firemní, zejména obchodní procesy a procesy řízení. Proces je definován jako množina jedné nebo více propojených činností společně přispívajících k dosažení vytyčeného cíle.

Řízení workflow zajišťuje Workflow management system, který definuje, vytváří a řídí průběh procesu, je schopen komunikovat s účastníky workflow a v případě potřeby spustit další aplikace. Workflow management system využívá současných komunikačních prostředků pro přenos potřebných informací a zvyšuje tak efektivnost celého procesu.

Při užívání workflow systému jsme schopni kdykoliv dohledat, v jaké fázi zpracování se dokument nachází, kdo jej vytvořil, kdo jej revidoval a kdo jej schválil, zajišťuje kontrolu nad dokumenty.

2.8 Bezpečnost (security)

V rámci CMS jsou definovány dva základní principy ochrany dokumentů. Zabezpečení na základě definování oprávnění pro určité uživatelské role a definování oprávnění pro určité dokumenty. Většina dnešních CMS umožňuje obě bezpečnostní strategie kombinovat.

V případě *oprávnění pro roli administrátor* přiřadí správce systému jednotlivým uživatelům nebo skupinám uživatelů jejich roli (např. tvůrce), tento uživatel pak může využívat pouze funkce této role (např. tvůrce může dokument vytvořit a zpracovat, ale nemůže jej následně revidovat).

Oprávnění pro dokumenty znamená definování přístupových práv pro daný archiv či dokument pro jednotlivé uživatele systému nebo skupiny uživatelů.

3. Typy CMS

V následující části je pozornost věnována jednotlivým typům CMS a jejich charakteristikám.

3.1 Systémy pro správu dokumentů (Document Management Systems)

DMS jsou komplexní systémy spravující nejrůznější typy dokumentů: texty, tabulky, prezentace, grafy, elektronické formuláře, faxy, skenované dokumenty, obrázky, fotografie, videozáznamy, zvukové soubory, technické výkresy i HTML stránky. Když má DMS otevřenou architekturu, která dovoluje sdílení SQL dat, pak mohou individuální aplikace sdílet data mezi sebou.

Typický DMS systém by měl obsahovat minimálně následující funkce:

- automatizovaný vstup papírových dokumentů (scanning/imaging) a organizace dokumentů do přehledné struktury (organizování dokumentů do elektronických pořadačů a spisů), umožňující rychlé vyhledání dokumentů podle zvolených kritérií,
- efektivní vyhledávání dokumentů (na základě klíčových slov i obsahu – fulltextový princip),

- volitelné atributy dokumentů – údaje blíže popisující daný dokument (metadata – např. autor, datum vzniku, název apod.),
- správa verzí dokumentu a audit – automatická tvorba a řízení verzí a revizí dokumentů (kdo, kdy a proč změnil dokument, s možností zachování jeho původní verze), řízení změn. Uživatelé mají jistotu, že pracují s poslední verzí dokumentů, mají také možnost prohlížet verze předchozí.
- automatická archivace, která zajišťuje, že po specifikované časové lhůtě jsou dokumenty smazány, přeneseny na archivační paměťové medium apod.,
- směrování a oběh dokumentů – nástroje pro realizaci přirozeného životního cyklu dokumentů, a to jak neřízeného, tak i řízeného, delegace procesů (ke konkrétnímu dokumentu lze definovat procesy různého typu); jde o řízenou editaci pro vybrané uživatele, schvalování faktur, objednávek, revize výkresové dokumentace, skartační mechanismy po uplynutí doby archivace, přezkoumání dokumentace ISO, oficiální notifikace o změně směrnic apod.,
- podpora elektronického schvalování a uvolňování dokumentů,
- publikace dokumentů na intranet – metody pro jednoduché zpřístupnění dokumentů jejich oprávněným odběratelům,
- zajištění bezpečnosti – přesné vymezení přístupových práv jednotlivých uživatelů dokumentů; nejčastější ohrožení dokumentů představují bývalí nebo stávající zaměstnanci organizace (např. lze zamezit kopírování a tisku),
- zvýšení efektivity práce uživatelů – možnost přizpůsobit klientskou stanici uživateli na míru (např. přidání tlačítek, předdefinované vzory dokumentů),
- tvorba reportů a monitoring – lze získat různé seznamy dle zadaných kritérií a sledovat kompletní činnosti prováděné s dokumentem, viz Lebeda (2004).

Důvodem pro nasazení DMS je tedy kvalitní archivace všech typů dokumentů, mnohonásobné urychlení vyhledávání konkrétního dokumentu, ochrana dat před možným zneužitím či neautorizovaným přístupem, okamžitá dostupnost aktuálních dat, finanční a časové úspory proti klasické údržbě archivu a zvýšení produktivity zaměstnanců.

3.2 Správa firemního obsahu (*Enterprise Content Management*)

Enterprise Content Management (dále ECM) je soubor strategií, nástrojů a technologií, používaných za účelem zachytit, spravovat, uložit, ochraňovat, distribuovat obsah a dokumenty v rámci organizačních procesů, bez ohledu na jejich formát [Asociace AIIM]. Neodmyslitelnou součástí ECM je už zmiňovaný DMS systém.

Za hlavní části ECM můžeme považovat:

- *document imaging* (DI) – převod do digitální formy,
- *document management* (DMS),
- *web content management* (WCM),
- *digital asset management* (DAM),
- *records management* (RM) – správa dokumentů, jejichž obsah již nelze měnit, a přesto musí být archivovány vzhledem ke své platnosti (např. podepsané smlouvy, přijaté faktury, účetní uzávěrky); jedná se více méně o elektronický archiv, který kontroluje skartační a archivační lhůty dokumentů,
- *team collaboration* (TCM) – slouží k podpoře rozhodovacích procesů, viz Tvrdíková (2002).

ECM pokrývá řízení a zpracování veškerého obsahu, který společnost vytváří a využívá. Za obsah se v tomto slova smyslu považují veškeré nestrukturované informace v rámci organizace. Nejdůležitějším předpokladem úspěšnosti každého ECM řešení je, aby obsah byl dostupný rychle a snadno všem, kdo s ním potřebují pracovat. Poté následují další důležité předpoklady, z nichž se k nejčastěji diskutovaným řadí bezpečnost a autenticita takto zpracovaných dokumentů.

3.3 Systémy pro zpracování elektronických formulářů (*Forms Processing*)

Oblast zpracování formulářů může být realizována zcela nezávisle nebo v rámci systémů CMS, DMS, ECM.

Systémy pro zpracování elektronických formulářů jsou schopny rozpoznat na základě definovaných vzorů daný formulář a vyčíst definované položky ve formuláři, lze definovat i kontroly na příslušné položky a tím zvýšit spolehlivost zpracování. Nabízí návrh sofistikovaných formulářů a jednoduché rozhraní pro uživatele. Formuláře mohou být propojeny na databáze podnikových aplikací. Některé systémy umožňují směřování formulářů dle předem stanovených pravidel a poskytují zpětnou informaci o stavu zpracování formuláře. Tyto aplikace umožňují pořízení digitálních dat v jednotné formě, rozsahu a obsahu, o jednotlivých případech téhož procesu za účelem dalšího jejich zpracování (pro vyhodnocení

procesů, zadání navazujících procesů, pořízení statistických údajů) viz Carda a Kunstová (2001). V případě potřeby může probíhat zpracování elektronických formulářů i paralelně, čímž se krátí čas zpracování.

Klíčem k vysoké úspěšnosti nasazení systému automatizovaného zpracování formulářů je komplexní analýza potřeb firmy či instituce a výběr a příprava jednotlivých komponent systému.

3.4 Workflow systémy

Architektura workflow systémů – model (podle společnosti *Workflow Management Coalition*) má následující strukturu:

- Programové komponenty modelu obsahují:
 - nástroj pro definici procesů umožňuje nadefinovat jednotlivé procesy, přiřadit role, stanovit pravidla (analýza a popis firemních procesů);
 - jádro workflow řídí průběh procesů, spouští externí aplikace, udržuje statistiky o průběhu workflow (procesní definice a organizační model);
 - správce úloh zajišťuje komunikaci mezi jádrem workflow a uživateli (nabídka položek k vyřízení, vyrovnávání pracovního vytížení uživatelů, znovu přidělování nevyřízených položek);
 - nástroj pro tvorbu jednotného uživatelského rozhraní (může tvořit se správcem úloh jeden celek).
- Datové komponenty modelu obsahují:
 - definice procesu, které popisují strukturu procesu (kroky, role, pravidla);
 - řídicí data workflow, jsou to interní data nutná pro zajištění chodu a znovuoobnovení procesu po havárii;
 - aplikační data (vlastní dokumenty), data externích aplikací;
 - pracovní data (data zpracovávaná jádrem systému), podle nich se rozhoduje o dalším kroku;
 - fronty úloh, datové struktury, ve kterých jsou uloženy všechny úkoly, které je třeba vykonat;
 - model organizační struktury, obsahuje data popisující organizační strukturu firmy, jednotlivé role, vztahy podřízenosti a nadřízenosti.

Workflow systém přispívá ke změně podnikových procesů, zlepšuje organizaci a kvalitu práce, zavedení standardních postupů zvyšuje efektivitu práce, pracovní postupy jsou uchovány v systému, ne v hlavách odcházejících pracovníků. Vyřizování případů se značně urychlí, v každém okamžiku je zjištělý stav

konkrétního případu, veškeré verze jsou autorizovány a jednotlivé verze uchovávány.

Automatizace workflow může být aplikována na širokém spektru procesů. Nejběžnějšími jsou: finance, účetnictví, marketing, skladové hospodářství, podpora zákazníků a prodej.

Významnou skupinu tvoří Coordination-based Workflow, která jsou zaměřena na komplexní automatizaci procesů. Kromě směřování dokumentů realizují i vazby na ostatní aplikace, viz Tvrđíková (2002).

3.5 Elektronická výměna dat (*Electronic Data Interchange*)

Další významnou technologii, která současně naplňuje i další cíle v usnadňování obchodu, je elektronická výměna dat (*Electronic Data Interchange* – EDI).

EDI je elektronická výměna strukturovaných standardních dat mezi dvěma aplikacemi dvou nezávislých subjektů, která je prováděna mezi systémy automaticky. Užitím této technologie lze posílat obchodní dokumenty přímo z interní počítačové aplikace do systému obchodního partnera, a to bez jakéhokoliv lidského zásahu. EDI umožňuje minimalizovat odpovědnost zaměstnanců a redukovat zpoždění a chyby, které ruční odbavování obchodních dokumentů provázejí. Vybírá informace z interních aplikací, překládá je do standardního formátu a odesílá v elektronické podobě po telefonních linkách či jiných telekomunikačních médiích. U příjemce jsou pak data znovu přeložena, a to přímo do interního informačního systému obchodního partnera.

EDI nejenže eliminuje papírovou korespondenci, jako je zasílání nákupních objednávek, faktur, transportních dokumentů, ale především umožňuje, aby veškeré transakce byly prováděny v reálném čase. EDI tak stojí v pozadí současného rozmachu just-in-time logistiky.

Aby ovšem automaticky zasílané informace nestály mimo výrobní operativu jednotlivých podniků, je nezbytností, aby EDI systémy byly plně integrovány s finančními, výrobními a distribučními moduly ERP systémů (*Enterprise Resource Planning*) výrobců a dodavatelů. V mnoha případech je však integrace ERP systémů s EDI komplikovaným a finančně náročným projektem. Je třeba dořešit systémová rozhraní, která by umožnila komunikaci těchto odlišných systémů. Nové, plně integrované EDI/ERP systémy však již umožňují, aby ERP data plynule proudila oběma směry – z a do systému, bez zvýšených nároků na údržbu systému jako celku.

EDI je elektronická výměna strukturovaných standardních zpráv mezi dvěma aplikacemi dvou nezávislých subjektů...

V této definici EDI: *Elektronická výměna dat* znamená výměnu dat elektronickou cestou (on line). Pro komunikaci lze využít sítě typu Internet, VDS Nextel, IBM IMNS, telefonní linky, rádiové nebo satelitní spojení. Elektronická výměna dat je většinou neinteraktivní, to znamená, že je nejprve vytvořen kompletní blok dat, který je odeslán najednou. Odesílání neprobíhá zároveň s tvorbou dat.

Výměna strukturovaných dat znamená, že data jsou definována syntaktickými pravidly. Tato pravidla vytvářejí vlastně společný jazyk pro všechny propojené aplikace. Syntaktická pravidla definují např. oddělovače nebo délky položek. Strukturovaná data jsou např. data v databázovém formátu, data pevné délky, data ve formátu CDF (Comma delimited format). Syntaxe strukturovaných zpráv je definována českou a mezinárodní normou ČSN ISO 9735 (UN/EDIFACT). Použití syntaxe je velmi důležité pro automatické zpracování. Použití mezinárodně platné normy je potřebné pro zajištění kompatibility s jinými EDI systémy.

Standardní zpráva je zpráva s předem definovaným typem, kde každá položka má své místo.

Část definice *mezi dvěma aplikacemi* upozorňuje na to, že EDI je určeno pro výměnu dat mezi dvěma aplikacemi na rozdíl od takových systémů, které jsou stavěny jako jednoúčelové aplikace (např. pro *homebanking*). Během zavádění systému se sice využívají zjednodušené aplikace, které mají jednoduchý interface pro ruční vkládání dat, ale konečným cílem je komunikace aplikace s aplikací.

Pasáž *mezi dvěma nezávislými subjekty* zdůrazňuje, že z hlediska návrhu je nutné definovat, zdali se jedná o návrh výměny dat mezi závislými, nebo nezávislými subjekty. Pokud má firma dvě kanceláře, každou na jiném konci města, nejedná se o nezávislé subjekty, ale o jediný podnik. Nemá smysl zavádět EDI pro vnitropodnikovou komunikaci. Pokud však např. výhradní odběratel zavádí EDI, je pro jeho dodavatele nutné se napojit, ale z koncepčního a především legislativního hlediska (právní stránka zde již hraje významnou roli) se jedná o nezávislé subjekty. Projekt je tedy nutné řešit jako EDI projekt. Výhradní odběratel by mohl definovat vlastní datové struktury, zabezpečení i komunikaci, ale je pro něj výhodnější držet se standardů, vzhledem k tomu, že v budoucnu bude nucen komunikovat s jiným (větším) subjektem – výhradní odběratel se státní správou, státní správa s Evropskou unií atd., a bude nucen používat jeho standardy. Je všestranně výhodnější rovnou vybudovat kvalitní systém, kde budou všichni účastníci užívat stejných standardů, než později celý systém měnit.

EDI musí splňovat tyto podmínky: *Integritu* – změna zprávy během přenosu bude odhalena, zpráva byla odeslána konkrétní osobou, zpráva přišla ve

správném pořadí; *Autentičnost* – určení osoby, která zprávu odeslala, neodmítnutí původu zprávy, neodmítnutí příjmu zprávy; *Důvěrnost zprávy* – zajištění obsahu zprávy před nepovolanými osobami; *Připuštění v právním rámci* – metody, které byly použity, jsou v souladu s právním kodexem.

V ČR dosud není vymezena konkrétní právní normou oblast elektronického obchodu. Toto vyžaduje, aby všechny zúčastněné strany o to pečlivěji smluvně vymezily vztahy mezi sebou navzájem.

Mezinárodním standardem pro elektronickou výměnu dat je systém *UN/EDIFACT (United Nations/Electronic Data Interchange For Administration, Commerce and Transport)*. Tento jediný mezinárodní standard pro elektronickou výměnu dat byl vypracován odbornými komisemi Organizace spojených národů a přijat v roce 1987. UN/EDIFACT je multiborová norma, která zastřešuje řadu podmnožin pro jednotlivá odvětví. V rámci jednotné úpravy mezinárodního standardu UN/EDIFACT jsou definovány stovky zpráv pro státní správu, obchod, dopravu, zdravotnictví, stavebnictví atp. Národním centrem pro UN/EDIFACT (norma ČSN ISO 9735) v České republice je FITPRO (*Facilitation of International Trade Procedures*) při Hospodářské komoře ČR. Vývoj standardů pro EDI probíhal zpočátku velmi nekoordinovaně. Postupem času docházelo ke vzniku odvětvových standardů na národní i mezinárodní úrovni. I dnes se tak proto vedle UN/EDIFACT používá několik dalších standardů, jako jsou SWIFT (bankovníctví), ODETTE (automobilový průmysl), ANSI X.12 (americká norma), SEDAS (rakouský systém), TRADACOM, UCS, TDCC, EDIA, WINS, VDA a další.

Rozsah užívání technologie EDI mezi obchodními partnery celosvětově neustále narůstá.

4. Práce s digitálními dokumenty

Následující část článku je věnována životnímu cyklu dokumentu a legislativě týkající se práce s dokumenty.

4.1 Životní cyklus listinných dokumentů v organizaci

Oběh dokumentů je upraven v zákoně o archivnictví a spisové službě 499/2004 Sb., vyhláška o podrobnostech výkonu spisové služby 646/2004 Sb. obsahuje tento tok dokumentů – příjem dokumentů, evidence dokumentů, rozdělování a oběh dokumentů, vyřizování dokumentů, vyhotovení dokumentů, podepisování dokumentů, odesílání dokumentů, ukládání dokumentů a archivnictví.

Vznik elektronického dokumentu

Dokument do organizace přichází či v organizaci vzniká v digitální formě nebo papírové formě.

V prvním případě je připraven pro zařazení do DMS systému, ve druhém případě se musí převést do elektronické podoby (skenování). Dokumenty přicházející do organizace se často již nemění, naopak dokumenty vzniklé uvnitř organizace se dále zpracovávají (mění se obsah a existuje několik verzí daného dokumentu).

Dokumenty vznikající uvnitř organizace lze dělit na výstupní dokumenty z firemních procesů (např. faktury a objednávky), ty se obvykle do DMS neukládají, interní dokumenty (určené pro vnitřní potřebu organizace, často důvěrné) a odchozí dokumenty (určené externím subjektům, např. smlouvy, nabídky).

Uložení do systému pro správu dokumentů

Před uložením dokumentů do úložiště dokumentů se dokument převede na vhodný formát a jsou mu přiřazena metadata (atributy).

Zpracování dokumentu

Dokumenty jsou uloženy v datovém úložišti a může se s nimi dále manipulovat. Ve většině případů je potřeba zajistit jejich doručení adresátovi a schválení nebo odeslání do dalších firemních procesů.

Pro zpracování dokumentů je důležité, aby systém umožňoval řízení práce více uživatelů nad jedním dokumentem, řízení jednotlivých verzí dokumentů, možnost elektronického schválení a ochranu dokumentů pomocí přístupových práv. K podpoře procesu zpracování dokumentů slouží workflow.

Archivace dokumentu

Po ukončení oběhu dokumentů v organizaci (dokument už nebude dále měněn) se rozhodne o jeho zařazení do archivu. Vybrané dokumenty se musí upravit do formy vhodné pro uložení, odeslání probíhá elektronicky nebo fyzickým doručením dokumentu na zvoleném médiu a pak jsou dokumentu přiřazeny indexy pro vyhledávání a identifikaci (dokument se většinou ukládá ve více kopiích pro případ selhání systému).

Při výběru je potřebné vycházet z legislativy, vnitřní politiky organizace a z požadavků archivu.

Pro zajištění čitelnosti archivovaných dokumentů je potřebné provádět kontrolu medií (média jsou testována na chyby při čtení uložených dat, v případě rozsáhlých archivů se uskutečňuje náátková kontrola dle plánu kontroly). Kontrolu formátů (rychlý vývoj IT má za následek rychle morální stárnutí použitých formátů v archivu): zastaralé formáty by měly být měněny na jiný formát, který zajistí jejich čitelnost do budoucna. Kontrolu zabezpečení (tak jako formáty, i použité metody dostupnosti dokumentů můžou postrádat účinnost díky rychlému vývoji IT): je tedy potřebné zastaralé mechanismy nahradit novým mechanismem. Kontrolu indexů, která ověřuje, že dokument lze pomocí indexů v archivu vyhledat.

Likvidace archivního balíku

Po skončení archivní lhůty jsou nepotřebné dokumenty z archivu vymazány spolu se všemi kopiemi. Tento krok musí být v souladu s platnou legislativou.

4.2 Legislativní omezení

Pro práci s dokumenty existuje celá řada legislativních omezení, zákonů a vyhlášek. Ze zákonů je to zákon č. 40/1964 Sb., občanský zákoník, zákon 235 2004 Sb. o DPH, zákon 101/2000 Sb. o ochraně osobních údajů a o změně některých zákonů, zákon 102/1971 Sb. o ochraně státního tajemství, zákon 106/1999 Sb. o svobodném přístupu k informacím, zákon 128/2000 Sb. o obcích, zákon 227/2000 Sb. o elektronickém podpisu, zákon 343/1992 Sb. o archivnictví, zákon 365/2000 Sb. o informačních systémech veřejné správy, zákon 190/2009 Sb. o archivnictví a spisové službě a o změně některých zákonů a zákon 29/2000 Sb. o poštovních službách.

Z vyhlášek pak vyhláška 191/2009 o podrobnostech výkonu spisové služby, vyhláška 117/1974 Sb., kterou se stanoví kritéria pro posuzování písemnosti jako archiválií a podrobnosti skartačního řízení, a vyhláška č. 366/2001 Sb. o upřesnění podmínek stanovených v § 6 a 17 zákona o elektronickém podpisu a o upřesnění požadavků na nástroje elektronického podpisu.

5. Návrh metodického rámce pro řešení přechodu na práci s digitálními dokumenty

Tento metodický rámec má podpořit návrh toku digitálních dokumentů v organizaci vytvářející nebo inovující svůj systém pro práci s dokumenty (pozornost je věnována také jejich archivaci). Jsou zde uváděny a charakterizovány jednotlivé technologické parametry pro práci s digitálními dokumenty a jejich archivaci, s cílem minimalizovat rizika spojená se specifickými vlastnostmi jednotlivých komponent vytvářeného systému.

5.1 Zpracování digitálních dokumentů

Jednotlivé fáze zpracování digitálních dokumentů jsou znázorněny na obrázku 2.

Příjem dokumentu

Listinné i digitální dokumenty jsou přijímány v podatelně. Dokumenty přijaté na jiném pracovišti se předávají do podatelny. Pracovník podatelny označí listinné dokumenty podacím razítkem a zaeviduje jejich přijetí ve spisové službě. Po zaevidování dokumentu jsou listinné dokumenty předány k převodu na digitální.

Převod na digitální dokument

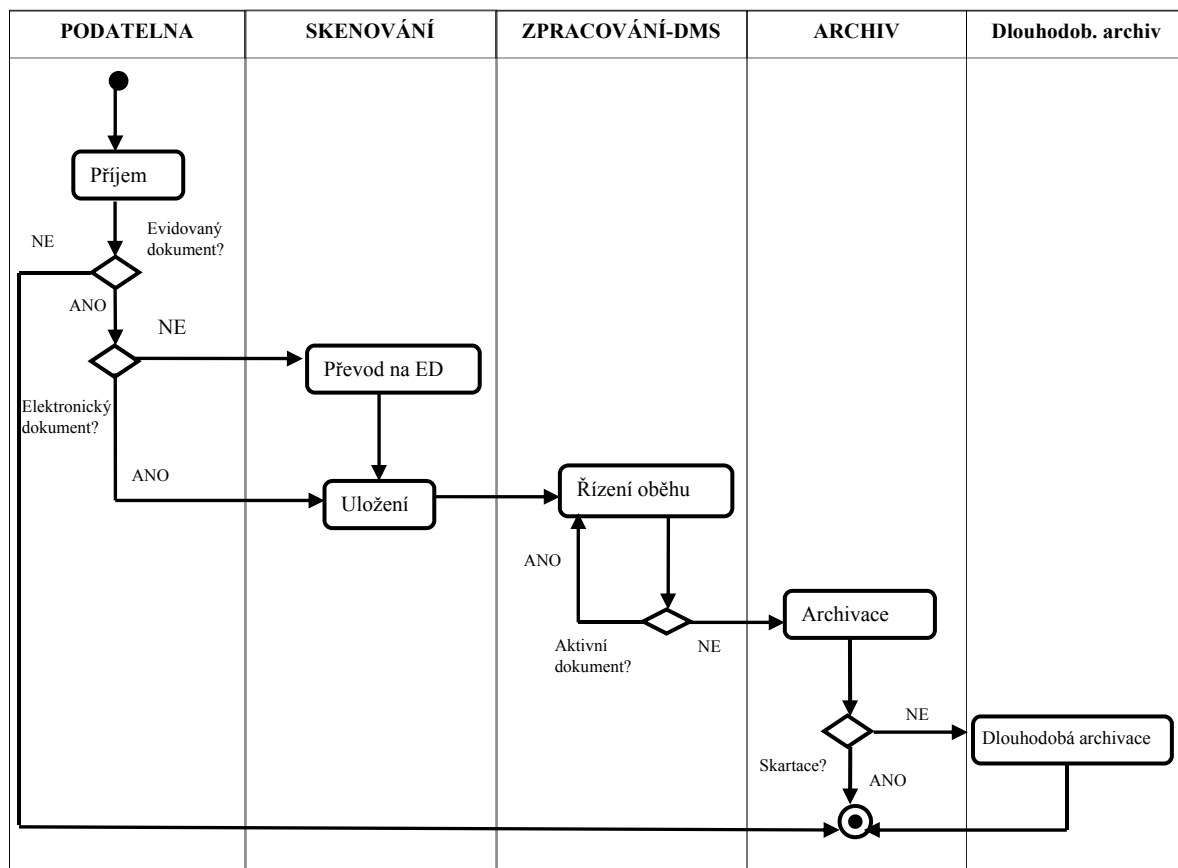
Převod listinného dokumentu na digitální dokument zahrnuje čtyři fáze:

- přípravu dokumentu, představuje třídění dokumentů podle typu; pro další zpracování dokumentu je nutné označit začátek nového dokumentu pomocí čárového kódu, který zároveň zajistí vazbu mezi elektronickou podobou dokumentu a originálem a může nést další informace, např. kategorii dokumentu,
- skenování a automatické zlepšení kvality naskenovaného dokumentu (dokument imaging), čímž se usnadní další krok, rozpoznávání,
- při rozpoznávání jsou informace z dokumentu převedeny do digitální podoby; současně pomocí technologie BCR dojde k vyhledání a přečtení čárového kódu. Informace z čárového kódu se uloží do metadat digitálního dokumentu,
- kontrolu a přiřazení metadat – před vlastním uložením nového digitálního dokumentu je nutné zkontrolovat kvalitu převedeného textu; kontrolu usnadní skenovací software označením těžko rozpoznatelných znaků a slov nenačtených ve slovníku.

Po kontrole a případné opravě chyb jsou v případě potřeby doplněna metadata. Metadata sestávají ze systémových atributů, generovaných systémem (např. datum, jméno uživatele, skenovací stanice, jednoznačný identifikátor dokumentu a archivu). Dále pak z atributů zadaných pracovníkem nebo získaných z dokumentu (nejčastěji čárový kód nebo atributy získané ze strukturovaného dokumentu pomocí technologie OCR) a atributů specifických pro daný dokument, které jsou také buďto zadávány obsluhou nebo doplněny systémem.

Ukládání

Při ukládání digitálního dokumentu je důležitý výběr vhodného formátu pro uložení. Při výběru formátu souboru pro archivační účely je prioritní zajištění přístupu k uloženým datům i v budoucnu, kdy nebude k dispozici stejné softwarové a hardwarové platformy, jak v době vzniku tohoto dokumentu. Proto se u archiválií kladě větší důraz na trvanlivost formátu než na snadnost použití nebo objemovou náročnost. V archivu se nepředpokládá častá frekvence přístupu k dokumentu, to neplatí pro dokumenty na začátku jejich životního cyklu, kdy je frekvence využívání (otevírání, editace, kopírování apod.) značná. Proto je v této fázi vhodnější použít běžné formáty. Nesplňují



Obrázek 2 Schéma toků digitálních dokumentů

všechny faktory pro dlouhodobou archivaci, ale vyznačují se lepší funkcionalitou. Jako jednoduchý příklad může posloužit nejrozšířenější formát textových dokumentů *.doc* programu *MS WORD*, který sice nesplňuje požadavky transparentnosti a otevřenosti formátu, ale nabízí uživateli větší komfort při práci než doporučený formát textových dokumentu pro dlouhodobou archivaci *.txt + TIFF*.

Řízení oběhu

Řízení oběhu dokumentů zajišťují systémy pro správu dokumentů, možnosti těchto systémů jsou různé, od základní evidence přijatých a odeslaných dokumentů ve spisové službě až po komplexní správu všech dokumentů během celého životního cyklu v systémech DMS.

Základní funkce systému správy dokumentů:

- Funkce správy, které usnadňují řízení, vytváření, úpravy, schvalování a používání elektronických dokumentů. Oběh dokumentu určují sami uživatelé výběrem dalšího zpracovatele (ad-hoc workflow) nebo na základě výběru jedné z přednastavených definic směřování (standardní workflow). Systém eviduje všechny události, které jsou významné pro historii dokumentu, zjednodušuje to vyhledávání doku-

mentu (vždy je možné zjistit, v jaké fázi životního cyklu se dokument nachází, osoby, které s dokumentem manipulovaly, dodržování termínů atp.).

- Funkce spolupráce, umožňující víceuživatelskou práci nad jedním dokumentem, řízení verzí, vytváření řešitelských týmů, které mají k dispozici stejné podklady, bez ohledu na jejich umístění.
- Funkce zabezpečení, definování přístupových práv k dokumentům, kontroly nad operacemi, které mohou uživatelé nad jednotlivými dokumenty provádět. Přístupová práva se mohou definovat pro skupiny uživatelů pomocí uživatelských rolí nebo detailním nastavením přístupových práv pro jednotlivé uživatele.
- Funkce zpřístupnění, zajišťující uživatelům vyhledání dokumentu, zobrazení jeho obsahu, editování a dělání kopií. Rozsah přístupových operací vychází z přístupových práv uživatele.

Archivace

Na konci oběhu dokumentu, za předpokladu, že už nebude měněn, nebo když se jedná o významnou verzi dokumentu, dojde k rozhodnutí, zda a jak bude dokument archivován. Při rozhodování se vychází z legisla-

tivy, ze zákona o archivnictví, zákona o spisové službě, politiky a procedur firmy, popřípadě z požadavků třetí strany.

Jestliže bude v organizaci většina (výhledově všechny) dokumentů v digitální formě, je nezbytné brát v úvahu také rizika jejich archivace. Jedná se o běžná rizika při jejich uchovávání a rizika hrozící po jejich uložení. Jako běžná rizika uchovávání digitálních dokumentů jsou chápány nefungující software, nefungující hardware, změna obsahu či smazání obsahu. Rizika, která ohrožují digitální dokumenty po uložení, jsou degradace nosiče, zastarávání HW, zastarávání formátu, zastarávání SW technologií a ztráta autenticity (platnosti autentizačních prvků).

Skartace

Dobu archivování dokumentů řídí zákon. Ten vymezuje skartační lhůty. Po ukončení skartační lhůty proběhne skartační řízení a dokumenty jsou trvale zničeny (dokumenty se skartačním znakem S), uloženy do archivu k trvalému uložení (dokumenty trvalé hodnoty se skartačním znakem A) nebo se posoudí, zda budou zničeny nebo dále archivovány (skartační znak V).

Dokumenty určené k trvalému uložení se přesunou do dlouhodobého archivu. Dlouhodobý archiv může mít firma či organizace vlastní nebo užívat archiv, který nabízí jiná organizace jako službu. Rozhodnutí závisí na množství dlouhodobě archivovaných dokumentů a schopnostech firmy zajistit podmínky pro dlouhodobou archivaci.

5.2 Technologické předpoklady vybudování účinného systému pro správu dokumentů

Výběr systému pro správu dokumentů a jejich archivaci je obtížný proces, protože prakticky neexistuje univerzální řešení, každý systém je potřeba upravit pro potřeby konkrétního nasazení. Prvním krokem při volbě systému je specifikace požadovaných vlastností systému. Požadované vlastnosti systémy jsou uváděny v tabulce 1.

Výběr softwarových komponent

Ústřední aplikací řešení je systém pro správu dokumentů starající se o celý životní cyklus elektronického dokumentu. Jako základ je vhodné použít komerční systém správy dokumentů a přizpůsobit jeho funkcionalitu požadavkům organizace úpravami nebo dodatečnými moduly. Doporučované možnosti výběru vhodného systému podle velikosti firmy jsou uváděny v tabulce 2.

Dalšími nezbytnými komponentami jsou výkonný a spolehlivý databázový systém a systém virtualizace datových úložišť poskytující prostředky pro jejich komplexní správu. K zajištění uživatelského rozhraní

Tabulka 1 Specifikace požadovaných vlastností systému

Požadavek	Zajišťovaná funkce	
Základní požadavky	Řízení změn, delegace procesů, reporty vyhledávání a monitoring	Tyto požadavky vycházejí z globální strategie a procesů probíhajících v organizaci
Spolehlivost a dostupnost	Umožnění přístupu k dokumentům po celou dobu jejich životního cyklu	Nezbytné pro všechny organizace
Bezpečnost	Řízení přístupu k dokumentům a zajištění integrity uložených dat	Nezbytné pro všechny organizace
Škálovatelnost	Možnost úprav systému v případě potřeby	Nezbytné pro všechny organizace
Snadná správa	Jednoduché řízení provozu systému	Nezbytné pro všechny organizace
Schopnost zaznamenávat historii	Evidování významných událostí v systému a operací prováděných s daty	Nezbytné pro všechny organizace

slouží aplikační a webové servery. Obrázek 3 znázorňuje architekturu systému.

Potřebné hardwarové komponenty a jejich funkce

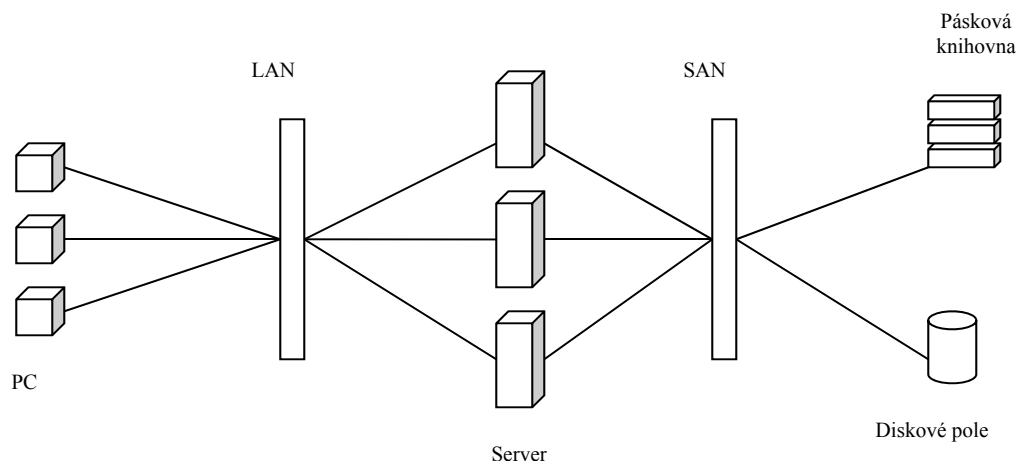
Pracovní stanice – PC jednotlivých pracovníků firmy.

Databázové servery – na databázových serverech pracují databáze obsahující metadata. Dostupnost a integrita databáze je klíčová pro chod celého systému.

Aplikační server – na aplikačním serveru pracuje aplikace pro správu dokumentů, která se stará o řízení životního cyklu elektronického dokumentu a poskytuje funkcionalitu potřebnou pro práci s dokumenty tak, aby uživatelé mohli pomocí vhodného rozhraní (nejlépe webového) vyhledat dokument, zobrazit, popřípadě získat kopii.

Diskové pole – diskové pole obsahuje databáze, je úložištěm vlastních dat. Podle strategie archivu budou na tomto úložišti uloženy všechny dokumenty nebo se dokumenty, u kterých je předpoklad, že nebudou v blízké době potřebné, přesunou na páskovou knihovnu. Diskové pole pro databáze by mělo být zaměřeno na výkon je u něj prioritní kapacita.

Pásková knihovna – slouží pro vytváření záloh a archivaci již neaktivních dokumentů. Za neaktivní dokument se považuje dokument, který již nebude dále měněn a je malá pravděpodobnost častého výskytu požadavku na jeho vyhledání.

**Obrázek 3** Architektura systému

Zdroj: Hejník, 2009

San (Storage Area Network) – datová dedikovaná síť pro paměťová zařízení, která slouží pro připojení externích zařízení k serverům, umožní sdílení zdrojů a vyšší propustnost.

Metadata

Metadata jsou v archivnictví velmi důležitým aspektem (nejznámějším představitelem metadat v klasickém archivnictví je katalogový lístek knihy, obsahující údaj o obsahu a uložení knihy v knihovně).

- Dublin Core – obsahuje sadu šestnácti prvků pro popis digitálních dokumentů. Často se používá jako základ pro metadata v archivech,

kde jsou základní prvky rozšířeny o další prvky. Dublin Core je standard pro metadatový popis digitálních objektů (včetně www stránek). Pracuje se na vytvoření české verze metadatového standardu Dublin Core pro popis a podporu vyhledávání elektronických informačních zdrojů v českém prostředí.

- OAIS – obecně popisuje entity a procesy, které by mělo repository (digitální úložiště) zahrnovat a dělat. Definuje také informační model pro metadata.

Tabulka 2 Doporučení pro výběr vhodného softwaru pro správu dokumentů

Typ CMS	Mikro firma nebo malá firma		Střední firma		Velká firma	
Dokument Imaging System	ano	nejvhodnější	ano	v rámci DMS	ano	v rámci ECM
Forms Processing	ano	podle charakteru činnosti firmy použitelné u firem různých velikostí	ano	dle charakteru činnosti	ano	v rámci ECM
Digital Asset Management	ano	pro speciálně zaměřené firmy na mediální dokumenty	ano	zaměřené na mediální dokumenty	?	dle charakteru činnosti
Web Content Management	ano	pro firmy, které se chtějí patřičně zviditelnit	ano	zviditelnění se firmy	ano	
Groupware System (komplexní)	ano	zejména pro geograficky rozptýlené firmy, podpora týmové práce	ano	v rámci DMS	ano	v rámci ECM
Dokument Management System	?	příliš nákladné	ano	nejvhodnější	ano	v rámci ECM
Workflow System	ne	příliš nákladné	ano	v rámci DMS	ano	v rámci ECM
Enterprise Content Management	ne	příliš nákladné	?	nákladné	ano	nejvhodnější
Electronic Data Interchange	ne	příliš nákladné	ano	vyžaduje-li to obchodní partner	ano	firmy s mnoha dodavateli či odběrateli (podpora Just in time logistiky)

Poznámka: Při členění firem na mikro firmy, malé firmy, střední firmy a velké firmy autor vycházel z počtu zaměstnanců ve firmě, tak jak jsou členěny v EU, avšak bez ohledu na velikost obrátu a velikost kapitálu nebo zisku.

Model OAIS byl přijat za normu (ISO 14721:2003).

- OAIS obecně popisuje entity a procesy, které by měl repositář zahrnovat a dělat, s tím, že určité části tohoto rámce mohou ve skutečném repositáři chybět. Repositář odpovídající OAIS pracuje s několika typy informačních balíčků. Jsou to tzv. SIP (Submission Information Package), AIP (Archival Information Package) a DIP (Dissemination Information Package). Ty mohou obsahovat informace o obsahu a samozřejmě k nim informace podpůrné, tj. konkrétní metadata, která jsou pro určitý balíček potřebná. Metadata informačních balíčků se od sebe liší a je v podstatě na instituci, jak si jejich podobu či strukturu nadefinuje.

Většina organizací, které se otázkou dlouhodobého uchovávání digitálních dokumentů seriálně zabývají, tento model více či méně implementuje. Všechny dostupné softwarové systémy pro správu úložišť ho také podporují, ať již jde o software komerční, nebo open source.

- PREMIS – popisuje podrobný návrh metadat pro digitální dokumenty, vychází z OAIS modelu a rozvádí jej do konkrétních detailů.

Popisná metadata – používají se pro vyhledávání a třídění, vychází ze standardu DUBLIN CORE.

Administrativní a technická metadata – metadata používaná pro správu elektronických dokumentů. Standard PREMIS.

Strukturální metadata – používají se pro strukturalizaci dokumentů (standard METS).

Kódování a uložení metadat

Výběr vhodného způsobu kódování metadat je stejně důležitý jako volba formátu souboru elektronického dokumentu. Kódování musí zajistit přístupnost a čitelnost metadat po celou dobu archivace vlastního dokumentu. Důležitým aspektem je možnost bezproblémového převodu metadat při přechodu archivu na novou technologii. Kódování musí vycházet z pevných standardů a být pokud možno nezávislé na použité technologii. V současnosti se jeví vhodným kódování s využitím standardů založených na XML. Často uváděným standardem v této oblasti je METS, který poskytuje kontejner, v rámci kterého mohou být uložena metadata z různých dalších standardů, které pak definují detailně sémantiku jednotlivých prvků. METS definuje pro XML šest základních částí pro umístění metadat (hlavička, popisná metadata, administrativní metadata, skupiny objektů, strukturální mapa a chování podle obsahu), viz Loeb, 2007.

Důležité je také vlastní uložení metadat. Lze využít tři základních možností:

- v relačních databázích,
- jako strukturovaná data uložená v XML nebo objektových databázích (tato technologie není prověřena pro velké objemy dat),
- jako XML soubory uložené společně s dokumenty – výhodou tohoto řešení je spojení metadat se souborem v jeden archivní celek, což zvyšuje důvěryhodnost a autentičnost.

Média pro ukládání digitálních dokumentů

Architektura úložiště dat může být založená pouze na užívání pevných disků. Přináší oproti tradičnímu zálohování na pásku výhodu v neustálé dispozici uložených dat, rychlých přístupových dobách, odpadá nutnost spravovat média. Ale i přes rostoucí kapacitu disků a jejich klesající ceny jsou disky pro dlouhodobou archivaci nevhodné. Důvodem jsou vyšší energetické nároky a krátký průměrný čas, který uplyne, než nastane chyba zařízení. Druhým řešením je zálohování na disk s archivací na magnetickou pásku. Toto řešení umožňuje, aby bylo dosaženo optimálního poměru mezi cenou a výkonem.

Při výběru vhodného média pro archivaci je rozhodující architektura systému úložiště dat a proces kontroly a údržby. Základními kritérii při výběru vhodného typu média jsou jeho životnost, kapacita, rychlost, odolnost a cena. Volit lze mezi pevnými disky (HDD), optickými disky a magnetickými páskami.

Chceme-li účinně bojovat se zastaráváním médií pro ukládání digitálních dokumentů, je nezbytné vytvořit vhodné podmínky pro jejich ukládání. Volit vhodný typ média, obměňovat média před dobou vypršení jejich životnosti. Vytvářet bezpečnostní kopie (kombinace médií) a provádět kontroly uložení (periodická detekce chyb).

6. Závěr

Charakteristickým rysem současné společnosti je nárůst transferu informací, lepší možnosti sdílení informací a umožnění všeobecného přístupu k veřejným informacím. Pro mnoho firem a organizací je dostupnost a rychlý pohyb informací rozhodující. V tabulce 3 je uveden přehled potřebných komponent CMS v závislosti na velikosti firmy.

Trh s CMS je široký. Při výběru je třeba vycházet z celkové informační strategie organizace a specifikovat cíle, kterých se má zavedením CMS dosáhnout, abychom neinvestovali neefektivně. Řízení a kontrola, kterou CMS přináší, transformuje soubor neorganizovaných dokumentů v dostupný informační majetek.

Tabulka 3 Přehled potřeby jednotlivých komponent *Content Management Systemu* podle velikosti firmy

Typ firmy Komponenta	Mikro a malá	Střední	Velká
Získávání	ano	ano	ano
Řízení	? Rozhoduje obor činnosti firmy	ano	ano
Skupinová spolupráce	ano	ano	ano
Ukládání	disk	dvouvrstvá technologie	dvouvrstvá technologie
Zabezpečení	ano	ano	ano
Doručení	nezáleží na velikosti, ale na oboru činnosti firmy		
Workflow	ne	? Rozhoduje obor činnosti firmy	ano
Bezpečnost	ano	ano	ano

Jestliže sledujeme oběh dokumentů v našich firmách a institucích, jde většinou o kombinaci digitálních záznamů a papírových dokumentů. Stav, kdy jsou papírové dokumenty přepisovány do počítače a pak opět vytisknuty, je neefektivní, přináší také velké riziko vzniku chyb a nepřesností.

Z výše uváděných důvodů lze konstatovat, že je dnes pro všechny firmy nutné zabývat se řešením problematiky práce s dokumenty, zvláště v návaznosti na povinnost užívání datových schránek.

Výhody digitálních dokumentů jsou zřejmé, je to úspora nákladů na tisk a doručení, úspora času a výhody plynoucí z jednodušší a levnější archivace dokumentů v digitální formě. Srovnání výhod práce s listinnými a digitálními dokumenty je v tabulce 4.

Tabulka 4 Porovnání zpracování listinných a digitálních dokumentů

	Listinné dokumenty ☒	Digitální dokumenty ☒
Náklady na pořízení	nízké	vyšší
Možnost ztráty	vysoká	minimalizovaná
Neproduktivní náklady (vyhledávání manipulace)	vysoké	minimalizovány
Dostupnost dokumentu	jediný uživatel	všichni, s povoleným přístupem
Ověřitelnost platnosti dokumentu	neověřitelný (existují neověřované kopie)	ověřitelný (digitální podpis)
Náklady na archivaci	vysoké (speciální skříně, místnosti, pracovníci archivu)	výrazně nižší (paměťová média, specialisté IT)

Další výhody užívání digitálních dokladů vyplývají z výhod elektronického podpisu, jedná se o obtížné falšování, ověřování integrity a nepopíratelnost.

Literatura

CARDA, A., KUNSTOVÁ, R. (2001). *Workflow, řízení firemních procesů*. Praha: Grada Publishing.

CARDA, A., KUNSTOVÁ, R. (2003). *Workflow. Nástroj manažera pro řízení podnikových procesů*. Praha: Grada Publishing.

HEJNÍK, J. (2009). *Návrh elektronické archivace dokumentů na městském úřadu*. Diplomová práce, Ostrava: VŠB-TU Ostrava.

LEBEDA, Z. (2004). Přínosy nasazení DMS systému. *IT systems* 1–2:33.

TĚTEK, M. (2006). Zamyšlení o content managementu. *IT Systems* 8: 12–16.

TVRDÍKOVÁ, M. (2008). *Aplikace moderních informačních technologií v řízení firmy – nástroje ke zvyšování kvality informačních systémů*. Praha: Grada Publishing.

TVRDÍKOVÁ, M. (2002). IT podpora správy dokumentů. *IT Systems* 11: 38–41.

TVRDÍKOVÁ, M. (2000). *Zavádění a inovace informačních systémů ve firmách*. Praha: Grada Publishing.

Další zdroje

LOEBL, Z. (2007) Dlouhodobé uchovávání elektronických dokumentů se zaručeným elektronickým podpisem. *Elektronická archivace* [Online]. [cit. 1. 4. 2008]. Dostupné z [www: <http://digiarchiv.eu/dokumenty/zprava/cast1.pdf>](http://digiarchiv.eu/dokumenty/zprava/cast1.pdf).

REINE, D. KAHN, M. (2008). Disk and Tape Square of Again: Tape Remains King of the Hill with LTO-4 2008th edition. Dostupné z [www: <http://www.cliper.com/publications.htm#Catalog>](http://www.cliper.com/publications.htm#Catalog).